

Dissecting hematopoietic and renal cell heterogeneity in adult zebrafish at single-cell resolution using RNA sequencing

Qin Tang,^{1,2,4,6*} Sowmya Iyer,^{1*} Riadh Lobbardi,^{1,2,4,6} John C. Moore,^{1,2,4,6} Huidong Chen,^{7,8,9} Caleb Lareau,^{3,7} Christine Hebert,^{1,10} McKenzie L. Shaw,^{1,10} Cyril Neftel,^{1,10} Mario L. Suva,^{1,10} Craig J. Ceol,¹¹ Andre Bernards,² Martin Aryee,^{3,7} Luca Pinello,^{1,2} Iain A. Drummond,⁵ and David M. Langenau^{1,2,4,6}

¹Molecular Pathology Unit, ²Center for Cancer Research, and ³Center for Computational and Integrative Biology, Massachusetts General Hospital, Charlestown, MA

⁴Center for Regenerative Medicine and ⁵Division of Nephrology, Department of Medicine, Massachusetts General Hospital, Boston, MA

⁶Harvard Stem Cell Institute, Harvard University, Cambridge, MA

⁷Department of Biostatistics, T.H. Chan School of Public Health, Harvard University, Boston, MA

⁸Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA

⁹Department of Computer Science and Technology, Tongji University, Shanghai, China

¹⁰Broad Institute, Cambridge, MA

¹¹Department of Molecular, Cell and Cancer Biology, University of Massachusetts Medical School, Worcester, MA

Recent advances in single-cell, transcriptomic profiling have provided unprecedented access to investigate cell heterogeneity during tissue and organ development. In this study, we used massively parallel, single-cell RNA sequencing to define cell heterogeneity within the zebrafish kidney marrow, constructing a comprehensive molecular atlas of definitive hematopoiesis and functionally distinct renal cells found in adult zebrafish. Because our method analyzed blood and kidney cells in an unbiased manner, our approach was useful in characterizing immune-cell deficiencies within DNA-protein kinase catalytic subunit (*prkdc*), interleukin-2 receptor γ (*il2rga*), and double-homozygous-mutant fish, identifying blood cell losses in T, B, and natural killer cells within specific genetic mutants. Our analysis also uncovered novel cell types, including two classes of natural killer immune cells, classically defined and erythroid-primed hematopoietic stem and progenitor cells, mucin-secreting kidney cells, and kidney stem/progenitor cells. In total, our work provides the first, comprehensive, single-cell, transcriptomic analysis of kidney and marrow cells in the adult zebrafish.

INTRODUCTION

A central challenge in developmental biology is to dissect the hierarchy of stem cells and to understand the functional diversity imparted to their differentiated progeny. Hematopoiesis is a beautiful example of how stem and progenitor cells give rise to mature cell types that nourish tissues by delivery of oxygen, protect the host from viral and bacterial pathogens, and provide clotting factors to stop blood cell loss after wounding (Copley et al., 2012). Zebrafish have emerged as a powerful, experimental tool for uncovering this functional heterogeneity, with fish sharing conserved cell types and underlying molecular programs that regulate hematopoiesis similar to those found in mammals (Dooley and Zon, 2000).

Despite clear similarities between zebrafish and mammalian blood cell development, definitive hematopoiesis occurs in a stromal compartment of the adult zebrafish kidney (Davidson and Zon, 2004), as opposed to the bone marrow in the mouse and human. Nonetheless, it is likely that hematopoietic stem and progenitor niche-supporting cells have conserved function and underlying molecular pathways responsible for driving continued blood production in adult fish. To date, the functional and transcriptional gene expression diversity of blood and their surrounding kidney cells has not been defined unbiasedly and at single-cell resolution in zebrafish.

Hematopoietic cell lineages can be commonly identified in mouse and human using cell-surface antibodies. However, that method has not been broadly applied to the zebrafish because of the lack of fish-specific antibodies. Rather, investigators have developed a wide array of fluorescent, transgenic, zebrafish lines that label specific blood cell types (Langenau et al., 2004; Mathias et al., 2006; Renshaw et al., 2006;

*Q. Tang and S. Iyer contributed equally to this paper.

Correspondence to David M. Langenau: dlangenau@mgh.harvard.edu

R. Lobbardi's present address is Blueprint Medicines, Cambridge, MA.

J.C. Moore's present address is Bluebird Bio, Cambridge, MA.

Abbreviations used: FISH, fluorescent in situ hybridization; GSEA, gene set enrichment analysis; HSPC, hematopoietic stem and progenitor cell; InDrops, indexing droplets single-cell RNA sequencing; lck, lymphocyte-specific protein tyrosine kinase; NK, NK-like; *prkdc*, DNA-protein kinase catalytic subunit; TALEN, transcription activator-like effector nuclease; Tg, transgenic; tSNE, *t* distributed stochastic neighbor embedding; ZFIN, zebrafish identifier nomenclature.

© 2017 Tang et al. This article is distributed under the terms of an Attribution-Noncommercial-Share Alike-No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms/>). After six months it is available under a Creative Commons License (Attribution-Noncommercial-Share Alike 4.0 International license, as described at <https://creativecommons.org/licenses/by-nc-sa/4.0/>).



Ellett et al., 2011; Page et al., 2013). Despite the great utility of fluorescent reporter lines for dynamic, in vivo cell imaging, those approaches often fail to identify the total heterogeneity of blood cell types because of the lack of sufficient transgenic lines and distinguishing fluorophores that are capable of labeling all the major blood cell types simultaneously within the same animal. Building on known transcriptional differences observed among blood cell types, we have previously developed a Fluidigm quantitative PCR platform using 96 PCR primers that can distinguish the major blood-cell lineages (Moore et al., 2016a). Developing that platform required prior knowledge of lineage-restricted marker genes, optimized PCR primer sets, and access to the expensive Fluidigm quantitative RT-PCR equipment. In contrast, Cvejic and colleagues have recently applied Smart-seq RNA, single-cell transcriptional profiling to identify subsets of $cd41^+$ hematopoietic stem and progenitor cells (HSPCs) and thrombocytes (Macaulay et al., 2016) and $lck:GFP^+$ lymphocytes (Carmona et al., 2017). However, to date, unbiased analysis of whole kidney marrow and large-scale RNA sequencing performed on thousands of individual cells from the same fish have not been reported, precluding unbiased identification of cell populations and quantitative assessment of how genetic mutations alter blood cell types within a given animal. In this study, we used droplet-based, single-cell RNA sequencing (Klein et al., 2015; Zilionis et al., 2017) to transcriptionally profile zebrafish blood and kidney cells. Using that approach, we identified novel blood and kidney cell types for which transgenic reporter lines are not yet available, discovered new marker genes for specified cell lineages in an unbiased manner, and characterized immune deficiencies within mutant zebrafish. Finally, we report a data set that can be queried to identify transcript differences within kidney marrow cells (Lareau et al., 2017). Those data will be useful in developing new transgenic reporter lines and in identifying novel markers of specific cell lineages.

RESULTS AND DISCUSSION

Single-cell RNA sequencing of blood cells from defined transgenic lines

In an effort to resolve hematopoietic cell lineages in adult zebrafish at single-cell resolution, we initially performed single-cell, transcriptomic profiling of cells isolated from fluorescent transgenic lines that label distinct blood-cell types. Specifically, whole kidney marrow and thymic T cells were isolated from adult, transgenic zebrafish using FACS, in which single cells were sorted into 96-well plates. RNA was isolated and made into bar-coded cDNA libraries using the Smart-seq2 method (Picelli et al., 2013). Transgenic lines analyzed by RNA sequencing included Tg(*Runx1+23:GFP*), which marks HSPCs (Tamplin et al., 2015); Tg(*cd41:GFP*), which labels both HSPCs and thrombocytes (Lin et al., 2005); Tg(*mpx:GFP*), which identifies neutrophil/myeloid lineages (Mathias et al., 2006; Renshaw et al., 2006); Tg(*rag2:GFP*), which marks marrow-derived B cells (Page et al., 2013); and Tg(*lck:GFP*),

which labels thymic T cells (Langenau et al., 2004). We also analyzed cells isolated from the whole kidney marrow of Tg(*lck:GFP*), *rag1*-deficient fish, which would be expected to lack mature T cells and yet retain GFP^+ NK cells (Shinkai et al., 1992; Wang et al., 1996; Moore et al., 2016a). In total, we profiled 246 single cells from those six transgenic lines, with a median of 1,059 genes detected per cell.

Differential gene-expression analysis across transgenic lines identified transcriptional programs consistent with identification of predicted blood-cell populations labeled within each transgenic line (Fig. 1 A). From that analysis, we identified known lineage-specific genes that were both highly expressed and found in a large fraction of isolated blood cells from each transgenic line (Fig. 1 B and Table S1). *t* distributed stochastic neighbor embedding (tSNE) analysis (van der Maaten and Hinton, 2008) was then used to visualize transcriptional differences among single cells. As expected, single cells from the same transgenic lineage largely clustered together. For example, tSNE visualization identified distinct clusters for *mpx:GFP^+* neutrophils, *lck:GFP^+* T and NK cells, and *rag2:GFP^+* B cells (Fig. 1 C). In contrast, cells from *cd41:GFP^{low}* and *Runx1+23:GFP* HSPCs failed to cluster into a single, defined group (Fig. 1 C), likely reflecting extensive and unpredicted heterogeneity within those transgenic, labeled cells.

Reconstruction of gene expression trajectories using Monocle and ARIADNE revealed expected delineation of T, NK, and B cells and neutrophil lineages (Fig. 1 D and Fig. S1 A). ARIADNE pseudo-time visualization also identified that *Runx1+23:GFP* and *cd41:GFP^{low}* cells were confined to two major trajectories; both of which expressed the stem and progenitor cell markers for T cell acute lymphocytic leukemia 1 (*tal1*), GATA binding protein 1a (*gata1a*), and v-myb avian myeloblastosis viral oncogene homologue (*myb*; Fig. 1 D and Fig. S1, B–E). One trajectory comprised classically defined HSPCs and expressed Meis homeobox 1 b (*meis1b*), pre-B cell leukemia homeobox 1a (*pbx1a*), zinc finger protein *zfp1* (*fog1*, friend of *gata1*), and friend leukemia integration 1 transcription factor (*flil1*; Fig. 1 D and Fig. S1). In contrast, a second trajectory characterized cells that had lineage skewing toward the erythroid lineage and expressed hemoglobin, α adult 1 (*hb α 1*), uroporphyrinogen decarboxylase (*urod*), spectrin β chain (*sptd*), and aminolevulinate, δ -, synthase 2 (*alas2*; Fig. 1 D and Fig. S1). These lineage assignments were confirmed using principle components analysis and unbiased hierarchical clustering using the top 20 differentially regulated genes identified in each ARIADNE trajectory (Fig. S1, B and C). Our results suggest a high degree of evolutionary conservation in blood cell hierarchies that are shared with mouse (Nestorowa et al., 2016) and human (Velten et al., 2017). Our results are also consistent with recent paradigm-shifting findings from humans that HSPCs can be lineage committed and that a subset of hematopoietic stem cells in adult marrow can directly specify erythroid cell types (Notta et al., 2016).

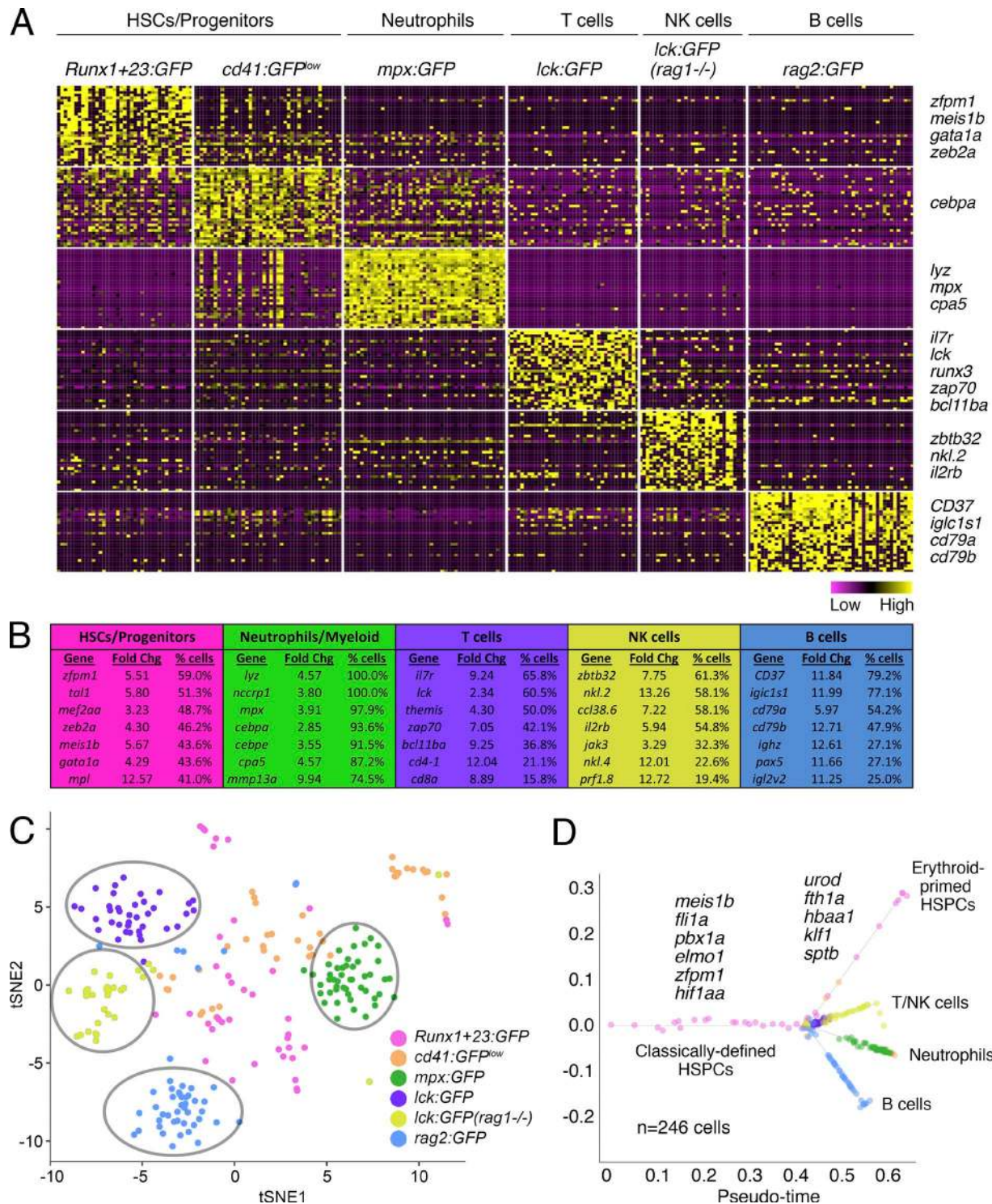


Figure 1. Single-cell RNA sequencing of blood cells isolated from transgenic zebrafish identifies novel gene signatures associated with specific hematopoietic cell lineages. (A) Heat map showing clustering with top 30 highly expressed genes within each transgenic line. Representative genes found within each specific cell lineage are denoted to the right. All cells were analyzed using the Smart-Seq2 protocol. (B) Expression of well-known, lineage-restricted genes expressed within fluorescent transgenic lines. Fold change (Chg) is represented as \log_2 -transformed values, and the percentage of cells (% cells) within a transgenic lineage that expresses a given gene is noted. (C) tSNE analysis showing population diversity of single cells sequenced from multiple fluorescent transgenic lines. Identified tSNE clusters are denoted by open circles. (D) Pseudo-time analysis

Unbiased analysis of whole-kidney marrow cells using InDrops RNA sequencing

Having defined many of the major hematopoietic cell lineages by sequencing highly purified, transgenic-labeled blood cells, we next profiled a larger number of unlabeled single cells isolated from the kidney of WT fish. These experiments sought to construct a cell lineage atlas of blood cells found within the adult zebrafish kidney marrow. To that end, we used a massively parallel transcriptomic method using indexing droplets (InDrops) single-cell RNA sequencing (Klein et al., 2015; Zilionis et al., 2017), which allowed high-throughput processing of thousands of cells from each individual kidney marrow. Specifically, single cells were harvested from the kidney of WT casper-strain zebrafish, reverse-transcribed, and uniquely bar-coded within individual droplets. cDNA from single cells was collectively processed for library construction and sequenced on a next-generation sequencing platform. Using this method, we profiled 3,782 cells from three WT, casper-strain animals, with a median of 811 genes detected per cell.

tSNE analysis was then used to unbiasedly identify molecularly defined clusters of marrow cells. Unbiased hierarchical clustering separated cells into the major hematopoietic cell lineages and also identified seven unique kidney stromal-cell types (Fig. 2, A and B). That analysis effectively made lineage assignments based on expression of well-known hematopoietic marker genes differentially enriched in each cluster. For example, high macrophage receptor with collagenous structure (*marco*) and macrophage expressed 1, tandem duplicate 2 (*mpeg1.2*) expression was indicative of the macrophage population, whereas cells in the erythroid cluster expressed hemoglobin subunit β -1 (*ba1*), *hbaa1*, and δ -aminolevulinic synthase 2 (*alas2*).

To confirm cluster-identity assignments, we also used gene signatures unbiasedly discovered from our SmartSeq2 single-cell RNA sequencing of fluorescent, transgenic hematopoietic lineages, projecting the combined expression of the top 20 most highly and frequently expressed genes found within each transgenically defined cell lineage (Table S2). From that analysis, we uncovered well-defined cell clusters that expressed signatures derived from *Runx1+23:GFP⁺* HSPCs, *lck:GFP⁺* T and NK cells, *rag2:GFP⁺* B cells, and *mpx:GFP⁺* neutrophils (Fig. 2, C and D; and Fig. S2). In analyzing genes identified in the overlap of the most highly and frequently expressed genes from the *Runx1+23:GFP* HSPCs (Table S2), we recognized that many of those genes included markers of the erythroid cell lineage, consistent with our SmartSeq2 RNA sequencing and pseudo-time trajectory analysis. Using the refined gene list identified from the trajectory analysis of classically defined HSPCs, we confirmed

that the same tSNE cluster was identified using this signature (Fig. S2, A–C). Additionally, we also analyzed a thrombocyte lineage signature generated from single-cell transcriptional profiling of *cd41:GFP^{high}* cells by Macaulay et al. (2016). Remarkably, that gene signature was also enriched within the classically defined HSPCs and the erythroid-primed HSPCs cluster (Fig. 2 D and Fig. S2 D). Those data raised the intriguing possibility that closely related transcriptional programs exist among classically defined HSPCs, erythroid lineage-committed progenitors, and thrombocytes and may be indicative of direct stem cell commitment into the megakaryocyte/erythroid cell lineage. Importantly, gene expression analysis of cell cycle-regulated genes confirmed that only a small fraction of HSPCs were cycling and identified further heterogeneity within the erythroid and neutrophil cells, likely identifying distinct progenitor and precursor cell states (Fig. S2 I; Tirosh et al., 2016).

Because *lck:GFP⁺* T cell and NK cell signature genes were both expressed within the same tSNE cluster (Fig. 2, C and D), we next performed additional tSNE analysis and visualization within that cell population, which revealed the existence of three molecularly distinct subclusters. Using gene expression signatures discovered from analysis of *lck:GFP⁺* cells from WT thymus or the marrow of *rag1*-deficient animals, we identified that the largest subcluster comprising 321 cells expressed T cell signature genes, which included CD8a molecule (*cd8a*), CD4-1 molecule (*cd4-1*), and ζ -chain (TCR)-associated protein kinase (*zap70*; Fig. 2 E and Table S3). Furthermore, a smaller subcluster comprised 142 cells and expressed both T and NK cell genes, including thymocyte selection associated (*themis*), T cell receptor β constant 1 (*trbc1*), and NK-lysin tandem duplicate 2 (*nkl.2*), and IL-2 receptor, β (*il2rb*; Fig. 2 F). Those data suggest that this subcluster comprised NK cells derived from the T cell lineage. Finally, we observed a cluster of 29 molecularly defined cells that failed to express either T or NK cell signature genes identified from sorted *lck:GFP⁺* cells, yet those cells expressed multiple chemokine (C-C motif) ligands (including *ccl33.3*), perforins (*prf1.2* and *1.7*), and NK-lysin 3 and 4 (*nkl.3* and *nkl.4*; Fig. 2, E and F; and Table S3). NK-lysin is analogous to mammalian granulysins, which are antimicrobial proteins made by both cytotoxic T lymphocytes and NK cells (Pereiro et al., 2015; Dotiwala et al., 2016). Those data suggest the existence of a second and novel NK-like cell population (NKL cells) that are not derived from *lck:GFP⁺* lymphocytes and yet retain genes expressed in cytotoxic and lytic cell types (Fig. 2 G). In total, our unbiased analysis of single-cell gene expression identified a wide array of expected and novel blood cell types within the marrow of the adult zebrafish.

sis showing predicted progression of cells through hematopoiesis. Differentially expressed genes indicative of classically defined and erythroid-primed HSPCs are demarcated on each trajectory. Sorted cells were combined from two independent, transgenic animals in each analysis ($n = 246$ total cells analyzed).

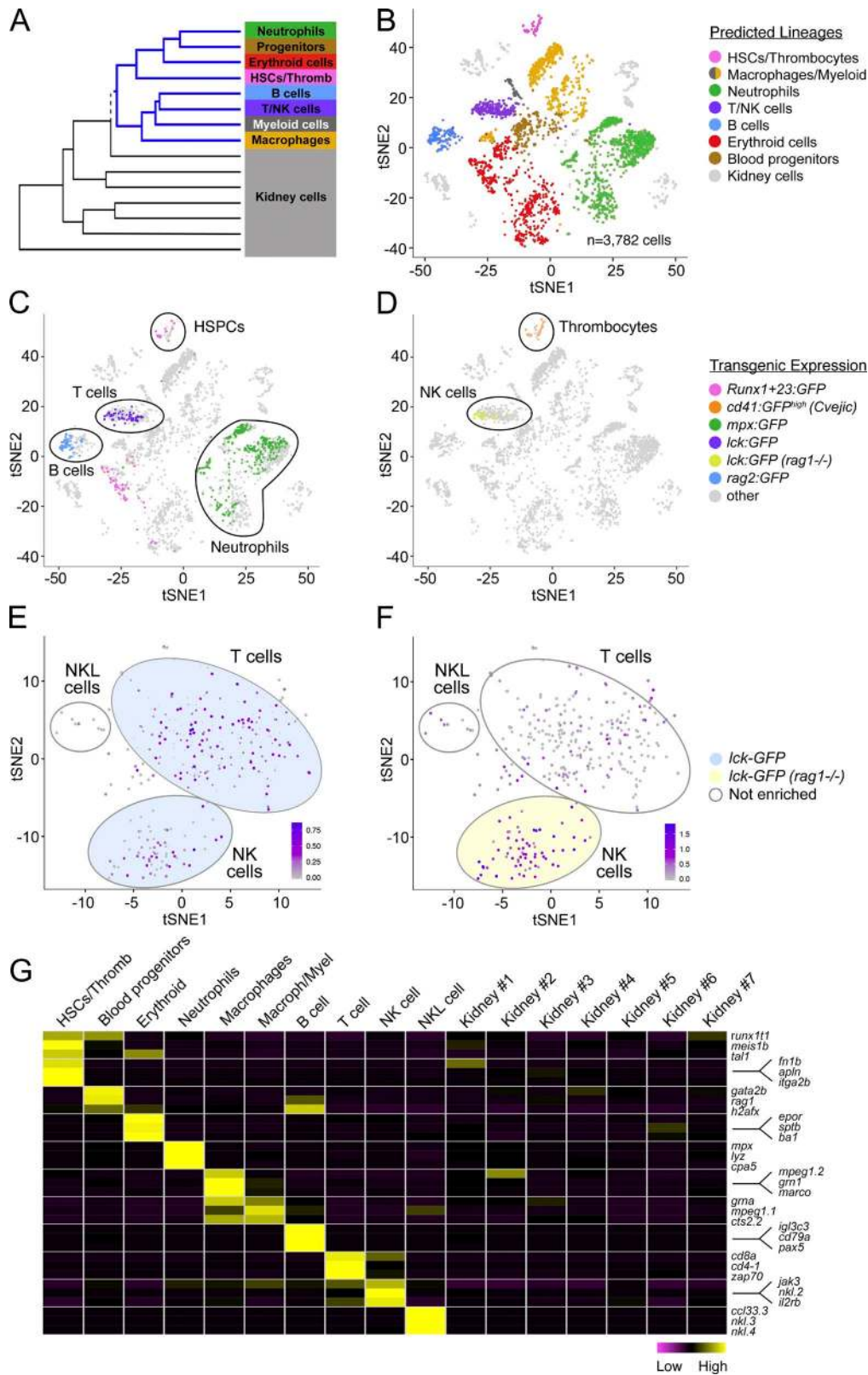


Figure 2. **Single-cell InDrops RNA sequencing of zebrafish whole kidney marrow.** (A) Hierarchical clustering identified cell lineages. Thromb, thrombocytes. (B) 2D projection of tSNE analysis. Hematopoietic cell lineages are denoted by colors. (C and D) Identification of hematopoietic cell types based on gene signatures identified from fluorescent transgenic zebrafish lines. Predicted cell types are denoted by open circles. (E and F) Three subgroups

Characterizing hematopoietic cell deficiencies in mutant zebrafish using InDrops sequencing

Next, we wanted to validate our proposed InDrops hematopoietic lineage identification by comparing single-cell transcriptional profiles isolated from marrow cells of WT and mutant animals. We have previously reported the characterization of DNA–protein kinase catalytic subunit (*prkdc*) homozygous mutant zebrafish (Moore et al., 2016b). These homozygous *prkdc*^{D36126} zebrafish have deficiencies in non-homologous end joining repair and thus fail to efficiently recombine T and B cell receptors, demonstrating striking diminution of B cells with only a modest reduction in T cell number when assessed by both quantitative real-time PCR analysis and RNA sequencing performed on bulk kidney marrow (Moore et al., 2016b). In this study, we profiled 3,201 single cells harvested from the kidney marrow of three homozygous *prkdc*^{D36126} mutant fish. We observed a 20-fold reduction in B cells in the homozygous *prkdc*^{D36126} mutant fish, whereas the percentage for T cells decreased by only one half (Fig. 3, A and B). *prkdc* deficiency specifically reduced the number of mature T cells and NK cells, whereas NKL cells were retained in homozygous *prkdc*^{D36126} mutant fish (Fig. 3, E and F).

To assess whether T and NK cell dysfunction could also be assessed using high-throughput single cell RNA sequencing methods, we created zebrafish with truncating mutations in the IL-2 receptor γ (*il2rga*) using transcription activator–like effector nuclease (TALEN)–mediated mutagenesis (Fig. S3). In mammals, the IL-2 receptor common γ chain dimerizes with a wide array of IL receptors to drive lymphocyte and NK cell maturation (Kondo et al., 1993, 1994; Russell et al., 1993; Matthews et al., 1995). Thus, deficiencies in this gene result in perturbed T and NK cell development. Indeed, histologic analysis of homozygous *il2rga*^{Y916} zebrafish revealed a dramatic loss of thymic T cells and a decrease in T and NK cell markers in the whole kidney marrow when assessed by quantitative PCR and bulk RNA sequencing (Fig. S3). As would be expected based on mouse and human deficiencies (Puck et al., 1997; Ito et al., 2002), B cells were unaffected in mutant fish (Fig. S3, D and E). Indeed, InDrops sequencing of homozygous *il2rga*^{Y916} mutant zebrafish also revealed a striking reduction in T and NK cell lineages with no overt reduction in B cells (Fig. 3, C and G; $n = 2,068$ single cells, two fish analyzed). In fact, the percentage of B cells increased relative to other hematopoietic groups in homozygous *il2rga*^{Y916} mutant fish, likely resulting from lineage compensation and shunting of lymphoid precursors into the B cell lineage. Lastly, generation of compound *prkdc*^{D36126}, *il2rga*^{Y916} double-homozygous mutant zebrafish resulted in losses in T, NK, and B cell populations (Fig. 3, D and H; $n = 2,276$ cells, two fish analyzed). In total, our experiments

provide a robust and efficient methodology to unbiasedly identify hematopoietic cell deficiencies in mutant zebrafish, a method likely to be useful for characterizing a wider array of mutant lines in the future.

Dissecting kidney cells at single-cell resolution

The vertebrate kidney has two main evolutionarily conserved functions. One is to remove waste substances from circulation, and the second is to balance osmolarity within a physiologic range (Vize et al., 1997). These functions are performed by highly conserved structures, including the glomerulus, segmented nephron tubules, and collecting duct (Vize et al., 1997). The simplicity of the kidney structure in the zebrafish embryo has propelled the model forward as one of the best for studying kidney development and function (Drummond et al., 1998; Morales and Wingert, 2017). Studies of the adult zebrafish kidney have also uncovered remarkable regenerative capacity and the identification of stem cell populations that generate new nephrons in response to injury (Diep et al., 2011). Remarkably, zebrafish hematopoiesis also takes place in the kidney, where blood cells are anatomically interspersed among a wide variety of kidney cell types; however, to date, molecularly defined kidney cells that maintain HSPC growth, self-renewal, and maturation have not been fully defined. Because we manually dissected whole kidney marrow, large numbers of kidney cells were also included in our single-cell InDrops sequencing, permitting a more-detailed characterization of those cell types in adult fish.

Our initial analysis of zebrafish kidney marrow samples identified distinct tSNE clusters that comprised seven different kidney cell types (Fig. 4 A). As may be expected, blood cell mutants had no overt differences in kidney cell composition when compared with those isolated from WT fish (Fig. 4 A, right). Similar to the analysis of hematopoietic cell lineages, we identified unique gene expression signatures that defined each of these kidney cell clusters ($n = 1,699$ total kidney cells analyzed; Table S3). The identities of four clusters could be readily assigned by expression of known genes. For example, three kidney clusters comprised epithelial cell types that are highly conserved with mammals (Lee et al., 2015), including (a) the proximal tubule defined by the expression of solute carrier family (*slc*) genes, (b) the distal tubule defined by expression of chloride channel K (*clcnk*), homeobox B8a (*hoxb8a*), and claudin h (*cldnh*), and (c) the nephron epithelial cells that share gene signatures with both proximal and distal tubules and may represent an intermediate tubule segment (Table S3). We also identified the long-postulated, yet molecularly undefined, mucin-secreting kidney cells. These cells expressed well-known markers of mucin-secreting cells, including oligomeric mucus/gel-forming (*muc2.4*), anterior

were identified within the T/NK cell population comprising T cells and NK cells that harbor *lck*:GFP signature, as well as NKL cells that fail to express *lck*:GFP signature genes. (G) Heat map denoting highly expressed genes indicative of each blood cell lineage. Data are rendered based on row normalization across samples. $n = 3,782$ cells analyzed from 3 WT fish.

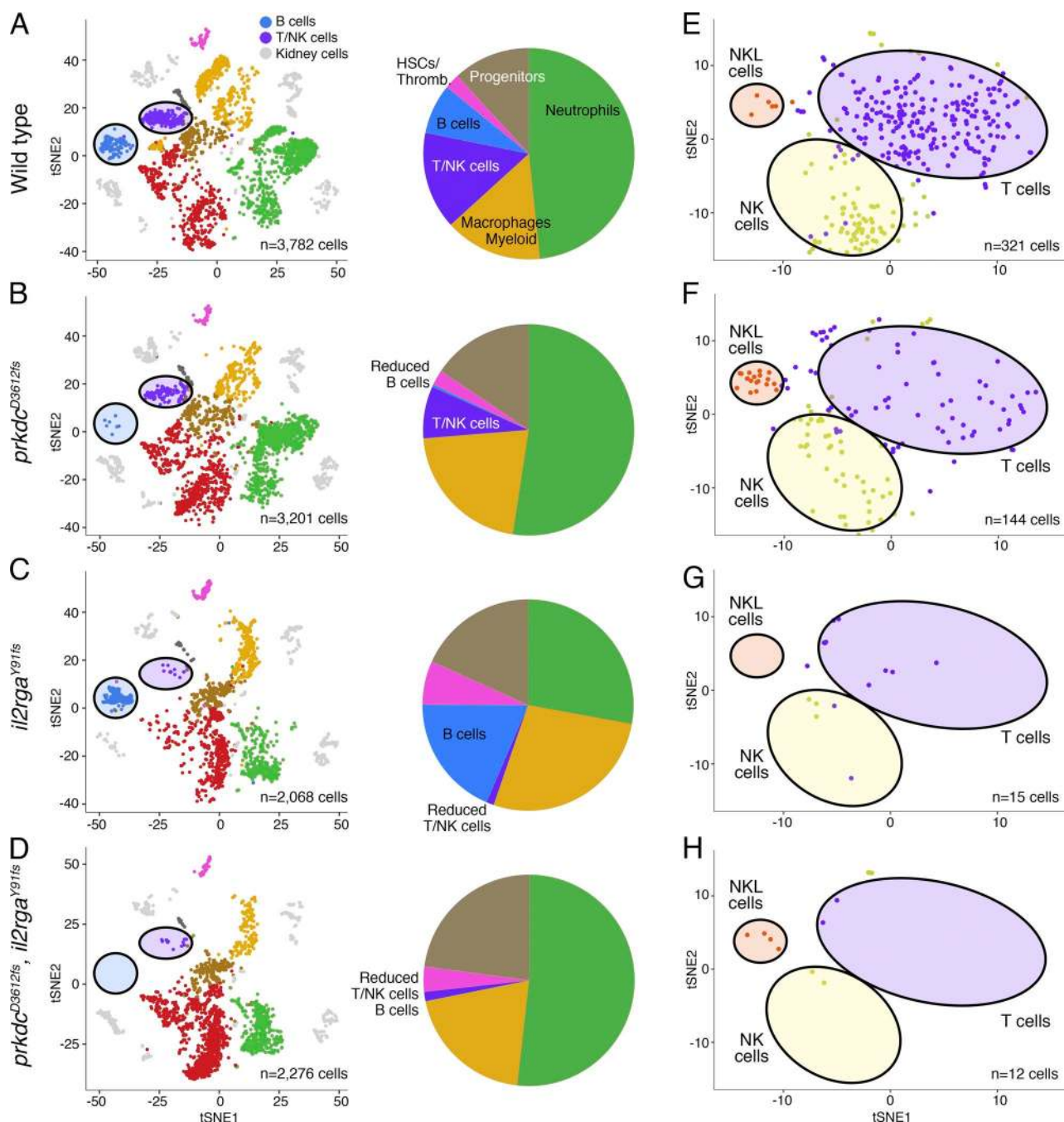


Figure 3. **Analysis of immunodeficient zebrafish using InDrops RNA sequencing of the whole kidney marrow.** (A–D) 2D projection of tSNE analysis for WT and mutant fish (left) and quantitation of white blood cells within each genotype of fish, demarcated as pie charts (right). (E–H) tSNE visualization showing T, NK, and NKL cell subpopulations within WT and mutant fish and denoted by shaded ovals. Number of cells within each analysis are noted. $n = 3$ animals for WT and *prkdc*^{-/-}; $n = 2$ for *il2rga*^{-/-} and double-mutant fish.

gradient 2 (*agr2*), and intelectin 1 (*itln1*; Fig. 4, B and C; and Table S3). These cells are likely responsible for secreting a mucin matrix to create immune barriers that curb infection, sharing a similar function to the mucous cells that line the epithelium of the mammalian urogenital tract (Brunskill and Potter, 2012). This is the first transcriptional analysis of this

unique cell type in fish, opening new opportunities to further study their biological functions and ontogeny with human kidney and other mucinous cell types.

The identities for the remaining three kidney cell clusters were identified using gene set enrichment analysis (GSEA), from which we uncovered prominent gene signa-

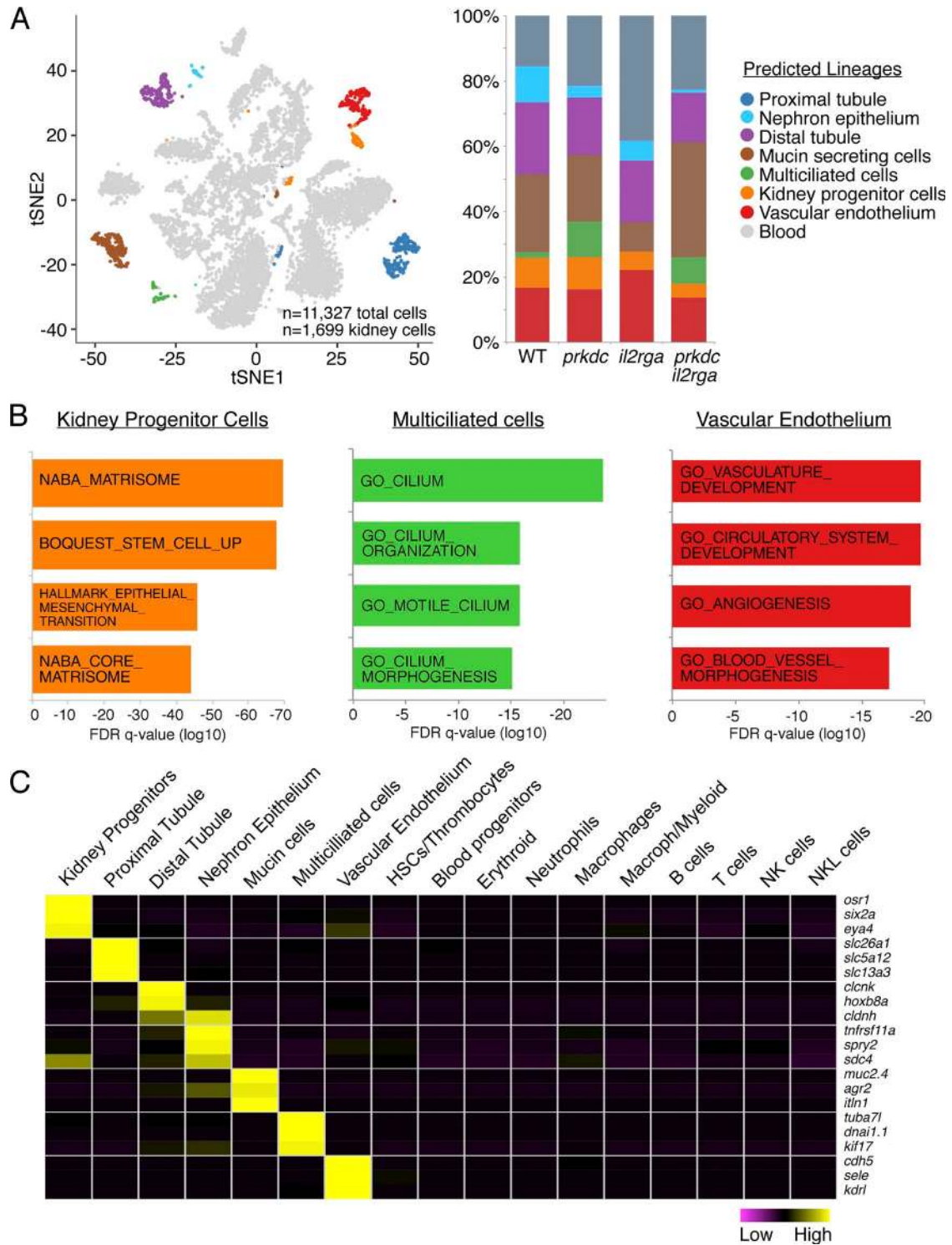


Figure 4. **Identification of novel renal cell lineages using InDrop RNA sequencing.** (A) 2D projection of tSNE analysis showing combined analysis of marrow cells from WT and mutant fish (left) and quantitation of identified kidney cell populations found within WT and mutant marrow (right). Kidney cell lineages denoted by colored dots, whereas blood cells are shown in gray. (B) Gene set enrichment analysis identified up-regulated pathways in kidney progenitor cells, multiciliated cells, and vascular endothelium. False-discovery rate p-values noted on the axis [false discovery rate (FDR) q-value]. (C) Heat map denoting highly expressed genes indicative of each kidney cell lineage. Data are rendered based on row normalization across samples. *n* = 10 fish from all combined genotypes.

tures for putative kidney stem/progenitor cells, multiciliated cells, and vascular endothelium (Fig. 4 B and Tables S4 and S5). All of the above identified kidney cell types also expressed the expected cell lineage-defining genes (Fig. 4 C). For example, the kidney progenitor cell population expressed known kidney stem/progenitor cell markers shared with the mouse metanephric cap mesenchyme (Davidson, 2008; Kobayashi et al., 2008), including SIX homeobox 2a (*six2a*), EYA transcriptional coactivator and phosphatase 2 (*eya2*), odd-skipped related transcription factor 1 (*osr1*), and transforming growth factor, β receptor II (*tgfbr2*). These kidney precursor cells were also enriched for genes highly expressed in human CD31⁻ stromal stem cells, yet also contained a large fraction of collagen matrix-associated genes (Fig. 4 B). Our analysis also identified the existence of molecularly distinct, multiciliated kidney cells that reside in the renal tubules to drive fluid movement through the kidney. These cells expressed well-known motor proteins required for cilia motility including tubulin α 7 like (*tuba7l*), dynein axonemal intermediate chain 1 paralog 1 (*dnai1.1*), and kinesin family member 17 (*kif17*; Fig. 4 C). Finally, GSEA MSigDB analysis allowed us to identify vascular endothelial cells that would be a predicted component of the kidney, supporting both healthy kidney cell function and blood cell trafficking from HSPC niches (Fig. 4, B and C). These cells expressed vascular endothelial growth factor receptor kdr-like (*kdr1*) and cadherin 5 (*cdh5*; vascular endothelial cadherin), with a subset of cells expressing *cxcl8a/8b*, a known chemokine required for sustaining HSPC maintenance and function (Blaser et al., 2017). In total, our work uncovered new kidney cell lineages not previously profiled transcriptionally and identified new marker genes for well-known cell types, opening new and exciting opportunities to study cell-type-specific roles in regulating kidney growth, regeneration, and niche support of hematopoietic stem cells.

Conclusion

Our work defined the single-cell heterogeneity of the zebrafish adult blood marrow and kidney cells at an unprecedented scale using both Smart-Seq2 and InDrops RNA sequencing approaches, providing a much-needed resource for the community to identify novel cell types and gene expression signatures that can be used as markers that define specific cell lineages in zebrafish. These data can be queried using a user-friendly website, defining InDrops gene expression within defined whole-kidney marrow cell clusters (see Lareau et al., 2017). When coupled with subcellular locational information using immunohistochemistry, fluorescent in situ hybridization (FISH), and/or transgenic approaches, the lineage-restricted genes identified by our studies will aid in spatially mapping the organization of blood and kidney cell types in the adult. This will be particularly useful in mapping vascular epithelial cell types that interact with HSPCs. Finally, as the cost of sequencing continues to drop and new methods are developed that allow transcriptional profiling at greater

depth within single cells, we envision that future studies will define blood and kidney cell heterogeneity at higher resolutions and likely allow assessment of intertwined relationships between zebrafish blood and kidney cells.

MATERIALS AND METHODS

Zebrafish

Transgenic reporter fish used in this work included Tg(*Runx1+23*:GFP; Tamplin et al., 2015), Tg(*cd41*:GFP; Lin et al., 2005), Tg(*mpx*:GFP; Mathias et al., 2006; Renshaw et al., 2006), Tg(*rag2*:GFP; Jessen et al., 2001), and Tg(*lck*:GFP; Langenau et al., 2004). Casper strain zebrafish (White et al., 2008), *prkdc*^{D36126} (ZFIN identifier, *prkdc*^{b103}) mutant fish (Moore et al., 2016b), and *rag1* mutant fish (Wienholds et al., 2002) have been reported previously. *il2rga*^{Y916} mutants were created in casper-strain zebrafish (ZFIN identifier, *il2rga*^{b104}) using TALEN-mediated mutagenesis, as previously described (Moore et al., 2012). TALEN targeting sequences were 5'-TACATTGAAAACAAGCCT-3' and 5'-AGTTTG GTGACGGAACAGGA-3'. *prkdc*^{D36126/D36126}, *il2rga*^{Y916/+} mutant fish were in-crossed to produce double-homozygous animals, which were identified by scale genotyping at 2–3 mo old (Moore et al., 2016b) and were generated at the expected Mendelian ratios. All zebrafish experiments were approved by the Massachusetts General Hospital Institutional Animal Care and Use Committee under protocol 2011N000127.

Isolation of single cells from the zebrafish kidney marrow

Zebrafish kidney marrow and thymus were dissected from 3–9-mo-old transgenic fish, made into single-cell suspensions, and individual cells were FACS sorted into 96-well plates, as previously described (Moore et al., 2016a). Each well contained 5 μ l TCL buffer (QIAGEN) + 1% β -mercaptoethanol (Bio-Rad Laboratories). 96-well plates were centrifuged at 1,000 g for 3 min at 4°C, snap-frozen on dry ice, and stored at -80°C. For InDrops sequencing approaches, marrow was isolated from zebrafish as single-cell suspensions and used in microfluidic sorting.

Smart-Seq2 single-cell RNA sequencing and processing

Whole-transcriptome amplification and library construction were performed on single cells using a modified Smart-Seq2 (Illumina) protocol adapted by (Venteicher et al., 2017). Libraries of uniquely bar-coded cells were combined and sequenced with a Nextseq 500 High Output V2 kit (75 cycles; Illumina) on a NextSeq 500 platform (Illumina). Using this method, we profiled $n = 246$ single cells from six transgenic lines.

The STAR program (Dobin et al., 2013) was used to align reads to the GRCz10 genome and Ensembl genome browser (version 85; EMBL-EML) was used for gene annotation. PCR duplicates were removed using Picard (Broad Institute), and reads were aligning to ribosomal RNA were removed using RSeqC split_bam.py. We then used featureCounts (Liao et al., 2014) to assign reads with alignment quality of 10 or more to

genes and generated a matrix of fragment counts in each cell. All further processing was performed using Seurat (GitHub). We filtered for cells that expressed a minimum of 500 genes and for genes that were expressed in ≥ 1 cell, with a gene detection threshold of five reads. To find markers of each transgenic population of cells, we combined expression read counts in each population and queried the differential expression of each gene in a foreground transgenic population versus all other cells. We used two metrics to calculate differential expression: (1) log-fold change of expression, and (2) difference in the percentage of cells expressing the gene in the foreground population versus the background population. Markers of *cd41*:GFP^{high} cells were identified as the top 20 genes ranked by correlation with high GFP expression from the data set produced by Macaulay et al. (2016). Human orthologues of zebrafish genes were found using the Beagle database (Bernards, 2017).

Indexing droplets (InDrops) sequencing and analysis

Microfluidic sorting and library construction were performed by the Harvard ICCB Single Cell Core using the method outlined in Klein et al. (2015) and Zilionis et al. (2017). In brief, kidney cell suspensions were resuspended in $1.0 \times \text{PBS} + 1\% \text{BSA}$ at $0.2\text{--}1 \times 10^6$ cells/ml, treated with 15% Optiprep (Sigma-Aldrich), and loaded onto the microfluidic sorting device. During microfluidic sorting, individual cells were captured in droplets, lysed, and RNA reverse transcribed using unique bar-codes, allowing measurement of the absolute transcript level by unique molecular identifier quantification within individual cells (Klein et al., 2015). Doublet rate was estimated at 4%, similar to that reported previously (Klein et al., 2015).

RNA from 1,500 cells of each sample kidney were pooled and subsequently processed together. cDNA products were preamplified to produce a cDNA library compatible with the Illumina Nextseq platform. Libraries were quantified using the KAPA Library Quantification kit (Kapa Biosystems). Three sequencing runs were completed using the Nextseq 500 High Output V2 kit (75 cycles; Illumina) on a NextSeq 500 platform (Illumina).

We used the InDrops pipeline (GitHub), modified to process zebrafish immune-cell data for processing InDrop sequencing reads. Reads were aligned to GRCz10 (Ensembl version 85) transcriptome using Bowtie. All further processing was performed using Seurat. We filtered for cells that expressed a minimum of 200 genes and filtering for genes that were expressed in a minimum of 10 cells, using a gene-detection threshold of one read. We combined expression data from all experiments ($n = 11,327$ cells), unbiasedly clustered cells by the first 25 principal components, and then visualized the data using tSNE. This allowed us to ensure that tSNE plots were uniform across experimental conditions and permitted us to digitally subtract samples to garner the renderings presented throughout this article. Further characterization the T/NK cell cluster identified three subclusters. Markers of each subcluster were identified using the same method described here and are shown in Table S3.

Cell cycle profiles were defined by first calculating the geometric mean expression of G1/S and G2/M genes identified by Tirosh et al. (2016) and then labeling the cell state as G1/S or G2/M if the geometric means of the respective genes in each phase were in the 95th percentile or higher.

RNA-seq on bulk zebrafish kidney marrow

Whole kidney marrow from WT, *prkdc*^{D36126}, *il2rga*^{Y916}, and double-mutant animals ($n = 3$ for each genotype) were surgically dissected and placed into 350 μl of QIAGEN RLT buffer containing 1% 2-mercaptoethanol, followed by RNA isolation using the RNAeasy Mini kit (QIAGEN). 100 ng of RNA isolated from each animal was subjected to mRNA selection by the RNA NEBNext Poly(A) mRNA magnetic isolation module (New England Biolabs, Inc.), followed by next-generation sequencing library construction using NEBNext Ultra Directional RNA Library Prep kit (New England Biolabs, Inc.) for Illumina using 15 cycles of PCR amplification. Libraries were sequenced on the Illumina HiSeq 2000 platform by the high-output, paired-end, 50-bp method with 1% PhiX spike-in.

The STAR program (Dobin et al., 2013) was used to align reads to GRCz10 genome and Ensembl (version 85) was used for gene annotation. PCR duplicates were removed using Picard, and reads aligning to ribosomal RNA were removed using RSeqC split_bam.py. We then used featureCounts (Liao et al., 2014) to assign reads with alignment quality of 10 or more to genes and generated a matrix of fragment counts in each cell. We used DESeq2 (Love et al., 2014) to detect differential expression between mutant and WT samples. We then identified genes that were significantly down-regulated [DESeq2 log₂(fc) less than -1 and adjusted $P < 0.05$] in each comparison. Human orthologues of each of these gene sets were obtained using Beagle and used for enrichment analysis with GSEA MSigDB.

Data deposition

Smart-Seq2 data for transgenic zebrafish lines, InDrops sequencing results, and RNA sequencing of total marrow from WT and mutant fish are available in the GEO database (accession nos. GSE100911, GSE100910, and GSE100912, respectively).

Pseudo-time analysis

To reconstruct cellular pseudo-time and to reorder cells into potential developmental trajectories, we used the ARIADNE software pipeline and Monocle2 obtained from Bioconductor (R package 'monocle' version 2.4.0; Trapnell et al., 2014; Qiu et al., 2017 *Preprint*). The ARIADNE method first represents single cells in a low-dimensional space based on transcriptomic similarities. Cellular trajectories are then modeled as a tree composed of curves inferred by principal curve fitting. After selecting a root state from the tree structure, ARIADNE generates a 2D representation with multiple branches, in which cells are plotted as dots, and cellular developmental trajectories are represented

as straight line. Along each trajectory, cells are ordered based on reconstructed pseudo-time. The lengths of trajectories and distances between cells and their projections on the trajectories are obtained from the low-dimensional space. PCA and hierarchical clustering were also completed on Tg(*Runx1+23*:GFP) cells using 40 differentially expressed genes automatically detected by ARIADNE (20 highly expressed only in the classically defined HSPCs trajectory and 20 highly expressed only in the erythroid-primed HSPCs trajectory).

For Monocle analysis, parameters were set to *expression-Family = 'tobit()'* and default parameters for the Smart-seq platform. For the comparison, we used the same gene-expression matrix used for the ARIADNE analysis.

Online supplemental information

Fig. S1 shows Monocle and ARIADNE pseudo-time analyses for transgenic-labeled blood cells profiled using SMA RTseq2. Fig. S2 shows the verification of InDrops cluster assignments using transgenically defined gene signatures. Fig. S3 presents the creation and characterization of the *il2rga*^{-/-} and *prkdc*^{-/-}, *il2rga*^{-/-} immune compromised zebrafish. Table S1 lists the genes denoted in the heat map provided in Fig. 1 A. Table S2 provides the Smart-Seq2 lineage signatures used for identification of InDrop groups in Fig. 2 (C–F). Table S3 provides InDrops RNA sequencing gene expression for each identified kidney and blood cell population. Table S4 lists the humanized genes used for GSEA signature analysis shown in Fig. 4 B. Table S5 provides an analysis of the differentially regulated genes by cluster using GSEAsig analysis. Table S6 provides the real-time quantitative PCR primers used in Fig. S3. Tables S1–S6 are provided as Excel files.

ACKNOWLEDGMENTS

We thank Dr. Felix Ellett, Lindsay Theodore, Dr. Finola Moore, and Elaine Garcia for animal husbandry and maintenance of transgenic zebrafish lines. We thank Christina C. Luo and Rachel Servis from Massachusetts General Hospital Flow Cytometry Core for help with single-cell sorting; Na Qu and Dr. Toshi Shioda for help with next-generation sequencing; Dr. Sarah A. Boswell, Alex Ratner, and Samuel L. Wolock from Harvard Medical School ICCB-L Single Cell Core for consultation on InDrops experimental design and technical assistance; Dr. Rory Kirchner for guidance on InDrops sequencing data analysis; and Dr. Ruslan Sadreyev, Anthony Anselmo, Ulandt Kim, and Yevgeniya Li from Massachusetts General Hospital Bioinformatics Core for their assistance with bulk RNA sequencing. We thank members of the Langenau laboratory for constructive discussion.

This work was supported by the National Institutes of Health, grant R24OD016761 to D.M. Langenau. Q. Tang was funded by the China Scholarship Council.

The authors declare no competing financial interests.

Author contributions: Q. Tang performed experiments, data analysis, and wrote the manuscript; S. Iyer performed bioinformatics analysis and wrote the manuscript; R. Lobbardi, J.C. Moore, C. Hebert, M.L. Shaw, and C. Neftel performed experiments, provided expert technical advice, and helped with experimental design; H. Chen and C. Lareau performed bioinformatics analysis and website design; A. Bernards helped with mapping zebrafish homologues with human genes and created the Beagle online portal; M.L. Suva, C.J. Ceol, M. Aryee, L. Pinello, I.A. Drummond, and D.M. Langenau provided help with experimental design, funding, project oversight, expert advice, and help with writing/editing the manuscript.

Submitted: 30 May 2017

Revised: 18 July 2017

Accepted: 26 July 2017

REFERENCES

- Bernards, A. 2017. The Beagle Genome Annotation Database. Harvard University. Available at: <http://chgr.mgh.harvard.edu/genomesrus/index.php>
- Blaser, B.W., J.L. Moore, E.J. Hagedorn, B. Li, R. Riquelme, A. Lichtig, S. Yang, Y. Zhou, O.J. Tamplin, V. Binder, and L.I. Zon. 2017. CXCR1 remodels the vascular niche to promote hematopoietic stem and progenitor cell engraftment. *J. Exp. Med.* 214:1011–1027. <http://dx.doi.org/10.1084/jem.20161616>
- Brunskill, E.W., and S.S. Potter. 2012. RNA-Seq defines novel genes, RNA processing patterns and enhancer maps for the early stages of nephrogenesis: Hox supergenes. *Dev. Biol.* 368:4–17. <http://dx.doi.org/10.1016/j.ydbio.2012.05.030>
- Carmona, S.J., S.A. Teichmann, L. Ferreira, I.C. Macaulay, M.J. Stubbington, A. Cvejic, and D. Gfeller. 2017. Single-cell transcriptome analysis of fish immune cells provides insight into the evolution of vertebrate immune cell types. *Genome Res.* 27:451–461. <http://dx.doi.org/10.1101/gr.207704.116>
- Copley, M.R., P.A. Beer, and C.J. Eaves. 2012. Hematopoietic stem cell heterogeneity takes center stage. *Cell Stem Cell.* 10:690–697. <http://dx.doi.org/10.1016/j.stem.2012.05.006>
- Davidson, A.J. 2008. Mouse kidney development. In *StemBook*. L. Girard, editor. Massachusetts General Hospital, Cambridge, MA. 1–19.
- Davidson, A.J., and L.I. Zon. 2004. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. *Oncogene.* 23:7233–7246. <http://dx.doi.org/10.1038/sj.onc.1207943>
- Diep, C.Q., D. Ma, R.C. Deo, T.M. Holm, R.W. Naylor, N. Arora, R.A. Wingert, F. Bollig, G. Djordjevic, B. Lichman, et al. 2011. Identification of adult nephron progenitors capable of kidney regeneration in zebrafish. *Nature.* 470:95–100. <http://dx.doi.org/10.1038/nature09669>
- Dobin, A., C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T.R. Gingeras. 2013. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics.* 29:15–21. <http://dx.doi.org/10.1093/bioinformatics/bts635>
- Dooley, K., and L.I. Zon. 2000. Zebrafish: A model system for the study of human disease. *Curr. Opin. Genet. Dev.* 10:252–256. [http://dx.doi.org/10.1016/S0959-437X\(00\)00074-5](http://dx.doi.org/10.1016/S0959-437X(00)00074-5)
- Dotiwala, F., S. Mulik, R.B. Polidoro, J.A. Ansara, B.A. Burleigh, M. Walch, R.T. Gazzinelli, and J. Lieberman. 2016. Killer lymphocytes use granulysin, perforin and granzymes to kill intracellular parasites. *Nat. Med.* 22:210–216. <http://dx.doi.org/10.1038/nm.4023>
- Drummond, I.A., A. Majumdar, H. Hentschel, M. Elger, L. Solnica-Krezel, A.F. Schier, S.C. Neuhaus, D.L. Stemple, F. Zwartkruis, Z. Rangini, et al. 1998. Early development of the zebrafish pronephros and analysis of mutations affecting pronephric function. *Development.* 125:4655–4667.
- Ellett, F., L. Pase, J.W. Hayman, A. Andrianopoulos, and G.J. Lieschke. 2011. mpeg1 promoter transgenes direct macrophage-lineage expression in zebrafish. *Blood.* 117:e49–e56. <http://dx.doi.org/10.1182/blood-2010-10-314120>
- Ito, M., H. Hiramatsu, K. Kobayashi, K. Suzue, M. Kawahata, K. Hioki, Y. Ueyama, Y. Koyanagi, K. Sugamura, K. Tsuji, et al. 2002. NOD/SCID/ γ_c^{null} mouse: An excellent recipient mouse model for engraftment of human cells. *Blood.* 100:3175–3182. <http://dx.doi.org/10.1182/blood-2001-12-0207>
- Jessen, J.R., T.N. Jessen, S.S. Vogel, and S. Lin. 2001. Concurrent expression of recombination activating genes 1 and 2 in zebrafish olfactory sensory neurons. *Genesis.* 29:156–162. <http://dx.doi.org/10.1002/gene.1019>

- Klein, A.M., L. Mazutis, I. Akartuna, N. Tallapragada, A. Veres, V. Li, L. Peshkin, D.A. Weitz, and M.W. Kirschner. 2015. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*. 161:1187–1201. <http://dx.doi.org/10.1016/j.cell.2015.04.044>
- Kobayashi, A., M.T. Valerius, J.W. Mugford, T.J. Carroll, M. Self, G. Oliver, and A.P. McMahon. 2008. Six2 defines and regulates a multipotent self-renewing nephron progenitor population throughout mammalian kidney development. *Cell Stem Cell*. 3:169–181. <http://dx.doi.org/10.1016/j.stem.2008.05.020>
- Kondo, M., T. Takeshita, N. Ishii, M. Nakamura, S. Watanabe, K. Arai, and K. Sugamura. 1993. Sharing of the interleukin-2 (IL-2) receptor gamma chain between receptors for IL-2 and IL-4. *Science*. 262:1874–1877. <http://dx.doi.org/10.1126/science.8266076>
- Kondo, M., T. Takeshita, M. Higuchi, M. Nakamura, T. Sudo, S. Nishikawa, and K. Sugamura. 1994. Functional participation of the IL-2 receptor γ chain in IL-7 receptor complexes. *Science*. 263:1453–1454. <http://dx.doi.org/10.1126/science.8128231>
- Langenau, D.M., A.A. Ferrando, D. Traver, J.L. Kutok, J.P. Hezel, J.P. Kanki, L.I. Zon, A.T. Look, and N.S. Trede. 2004. In vivo tracking of T cell development, ablation, and engraftment in transgenic zebrafish. *Proc. Natl. Acad. Sci. USA*. 101:7369–7374. <http://dx.doi.org/10.1073/pnas.0402248101>
- Lareau, C., S. Iyer, D.M. Langenau, and M. Aryee. 2017. Single Cell inDrops RNA-Seq Visualization of Adult Zebrafish Whole Kidney Marrow. Harvard University. Available at: <https://molpath.shinyapps.io/zebrafishblood/>
- Lee, J.W., C.L. Chou, and M.A. Knepper. 2015. Deep sequencing in microdissected renal tubules identifies nephron segment-specific transcriptomes. *J. Am. Soc. Nephrol*. 26:2669–2677. <http://dx.doi.org/10.1681/ASN.2014111067>
- Liao, Y., G.K. Smyth, and W. Shi. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 30:923–930. <http://dx.doi.org/10.1093/bioinformatics/btt656>
- Lin, H.F., D. Traver, H. Zhu, K. Dooley, B.H. Paw, L.I. Zon, and R.I. Handin. 2005. Analysis of thrombocyte development in CD41-GFP transgenic zebrafish. *Blood*. 106:3803–3810. <http://dx.doi.org/10.1182/blood-2005-01-0179>
- Love, M.I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 15:550. <http://dx.doi.org/10.1186/s13059-014-0550-8>
- Macauley, I.C., V. Svensson, C. Labalette, L. Ferreira, F. Hamey, T. Voet, S.A. Teichmann, and A. Cvejic. 2016. Single-Cell RNA-sequencing reveals a continuous spectrum of differentiation in hematopoietic cells. *Cell Reports*. 14:966–977. <http://dx.doi.org/10.1016/j.celrep.2015.12.082>
- Mathias, J.R., B.J. Perrin, T.X. Liu, J. Kanki, A.T. Look, and A. Huttenlocher. 2006. Resolution of inflammation by retrograde chemotaxis of neutrophils in transgenic zebrafish. *J. Leukoc. Biol*. 80:1281–1288. <http://dx.doi.org/10.1189/jlb.0506346>
- Matthews, D.J., P.A. Clark, J. Herbert, G. Morgan, R.J. Armitage, C. Kinnon, A. Minty, K.H. Grabstein, D. Caput, P. Ferrara, et al. 1995. Function of the interleukin-2 (IL-2) receptor γ -chain in biologic responses of X-linked severe combined immunodeficient B cells to IL-2, IL-4, IL-13, and IL-15. *Blood*. 85:38–42.
- Moore, F.E., D. Reyon, J.D. Sander, S.A. Martinez, J.S. Blackburn, C. Khayter, C.L. Ramirez, J.K. Joung, and D.M. Langenau. 2012. Improved somatic mutagenesis in zebrafish using transcription activator-like effector nucleases (TALENs). *PLoS One*. 7:e37877. <http://dx.doi.org/10.1371/journal.pone.0037877>
- Moore, F.E., E.G. Garcia, R. Lobbardi, E. Jain, Q. Tang, J.C. Moore, M. Cortes, A. Molodtsov, M. Kasheta, C.C. Luo, et al. 2016a. Single-cell transcriptional analysis of normal, aberrant, and malignant hematopoiesis in zebrafish. *J. Exp. Med*. 213:979–992. <http://dx.doi.org/10.1084/jem.20152013>
- Moore, J.C., Q. Tang, N.T. Yordán, F.E. Moore, E.G. Garcia, R. Lobbardi, A. Ramakrishnan, D.L. Marvin, A. Anselmo, R.I. Sadreyev, and D.M. Langenau. 2016b. Single-cell imaging of normal and malignant cell engraftment into optically clear prkdc-null SCID zebrafish. *J. Exp. Med*. 213:2575–2589. <http://dx.doi.org/10.1084/jem.20160378>
- Morales, E.E., and R.A. Wingert. 2017. Zebrafish as a model of kidney disease. *Results Probl. Cell Differ*. 60:55–75. http://dx.doi.org/10.1007/978-3-319-51436-9_3
- Nestorowa, S., F.K. Hamey, B. Pijuan Sala, E. Diamanti, M. Shepherd, E. Laurenti, N.K. Wilson, D.G. Kent, and B. Göttgens. 2016. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood*. 128:e20–e31. <http://dx.doi.org/10.1182/blood-2016-05-716480>
- Notta, F., S. Zandi, N. Takayama, S. Dobson, O.I. Gan, G. Wilson, K.B. Kaufmann, J. McLeod, E. Laurenti, C.F. Dunant, et al. 2016. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*. 351:aab2116. <http://dx.doi.org/10.1126/science.aab2116>
- Page, D.M., V. Wittamer, J.Y. Bertrand, K.L. Lewis, D.N. Pratt, N. Delgado, S.E. Schale, C. McGue, B.H. Jacobsen, A. Doty, et al. 2013. An evolutionarily conserved program of B-cell development and activation in zebrafish. *Blood*. 122:e1–e11. <http://dx.doi.org/10.1182/blood-2012-12-471029>
- Pereiro, P., M. Varela, P. Diaz-Rosales, A. Romero, S. Dios, A. Figueras, and B. Novoa. 2015. Zebrafish NK-lysins: First insights about their cellular and functional diversification. *Dev. Comp. Immunol*. 51:148–159. <http://dx.doi.org/10.1016/j.dci.2015.03.009>
- Picelli, S., A.K. Björklund, O.R. Faridani, S. Sagasser, G. Winberg, and R. Sandberg. 2013. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods*. 10:1096–1098. <http://dx.doi.org/10.1038/nmeth.2639>
- Puck, J.M., A.E. Pepper, P.S. Henthorn, F. Candotti, J. Isakov, T. Whitwam, M.E. Conley, R.E. Fischer, H.M. Rosenblatt, T.N. Small, and R.H. Buckley. 1997. Mutation analysis of IL2RG in human X-linked severe combined immunodeficiency. *Blood*. 89:1968–1977.
- Qiu, X., Q. Mao, Y. Tang, L. Wang, R. Chawla, H. Pliner, and C. Trapnell. 2017. Reversed graph embedding resolves complex single-cell developmental trajectories. *bioRxiv*. doi:10.1101/110668 (Preprint posted February 21, 2017).
- Renshaw, S.A., C.A. Loynes, D.M. Trushell, S. Elworthy, P.W. Ingham, and M.K. Whyte. 2006. A transgenic zebrafish model of neutrophilic inflammation. *Blood*. 108:3976–3978. <http://dx.doi.org/10.1182/blood-2006-05-024075>
- Russell, S.M., A.D. Keegan, N. Harada, Y. Nakamura, M. Noguchi, P. Leland, M.C. Friedmann, A. Miyajima, R.K. Puri, W.E. Paul, et al. 1993. Interleukin-2 receptor γ chain: A functional component of the interleukin-4 receptor. *Science*. 262:1880–1883. <http://dx.doi.org/10.1126/science.8266078>
- Shinkai, Y., G. Rathbun, K.P. Lam, E.M. Oltz, V. Stewart, M. Mendelsohn, J. Charron, M. Datta, F. Young, A.M. Stall, et al. 1992. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D) J rearrangement. *Cell*. 68:855–867. [http://dx.doi.org/10.1016/0092-8674\(92\)90029-C](http://dx.doi.org/10.1016/0092-8674(92)90029-C)
- Tamplin, O.J., E.M. Durand, L.A. Carr, S.J. Childs, E.J. Hagedorn, P. Li, A.D. Yzaguirre, N.A. Speck, and L.I. Zon. 2015. Hematopoietic stem cell arrival triggers dynamic remodeling of the perivascular niche. *Cell*. 160:241–252. <http://dx.doi.org/10.1016/j.cell.2014.12.032>
- Tirosh, I., B. Izar, S.M. Prakadan, M.H. Wadsworth II, D. Treacy, J.J. Trombetta, A. Rotem, C. Rodman, C. Lian, G. Murphy, et al. 2016. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. 352:189–196. <http://dx.doi.org/10.1126/science.aad0501>

- Trapnell, C., D. Cacchiarelli, J. Grimsby, P. Pokharel, S. Li, M. Morse, N.J. Lennon, K.J. Livak, T.S. Mikkelsen, and J.L. Rinn. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 32:381–386. <http://dx.doi.org/10.1038/nbt.2859>
- van der Maaten, L., and G. Hinton. 2008. Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9:2579–2605.
- Velten, L., S.F. Haas, S. Raffel, S. Blaszkiewicz, S. Islam, B.P. Hennig, C. Hirche, C. Lutz, E.C. Buss, D. Nowak, et al. 2017. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat. Cell Biol.* 19:271–281. <http://dx.doi.org/10.1038/ncb3493>
- Venteicher, A.S., I. Tirosh, C. Hebert, K. Yizhak, C. Neftel, M.G. Filbin, V. Hovestadt, L.E. Escalante, M.L. Shaw, C. Rodman, et al. 2017. Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science.* 355:eaai8478. <http://dx.doi.org/10.1126/science.aai8478>
- Vize, P.D., D.W. Seufert, T.J. Carroll, and J.B. Wallingford. 1997. Model systems for the study of kidney development: Use of the pronephros in the analysis of organ induction and patterning. *Dev. Biol.* 188:189–204. <http://dx.doi.org/10.1006/dbio.1997.8629>
- Wang, B., G.A. Hollander, A. Nichogiannopoulou, S.J. Simpson, J.S. Orange, J.C. Gutierrez-Ramos, S.J. Burakoff, C.A. Biron, and C. Terhorst. 1996. Natural killer cell development is blocked in the context of aberrant T lymphocyte ontogeny. *Int. Immunol.* 8:939–951. <http://dx.doi.org/10.1093/intimm/8.6.939>
- White, R.M., A. Sessa, C. Burke, T. Bowman, J. LeBlanc, C. Ceol, C. Bourque, M. Dovey, W. Goessling, C.E. Burns, and L.I. Zon. 2008. Transparent adult zebrafish as a tool for in vivo transplantation analysis. *Cell Stem Cell.* 2:183–189. <http://dx.doi.org/10.1016/j.stem.2007.11.002>
- Wienholds, E., S. Schulte-Merker, B. Walderich, and R.H. Plasterk. 2002. Target-selected inactivation of the zebrafish *rag1* gene. *Science.* 297:99–102. <http://dx.doi.org/10.1126/science.1071762>
- Zilionis, R., J. Nainys, A. Veres, V. Savova, D. Zemmour, A.M. Klein, and L. Mazutis. 2017. Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protoc.* 12:44–73. <http://dx.doi.org/10.1038/nprot.2016.154>